

Von LaTeX zu PDF

Elke Salecker
saleckee@cms.hu-berlin.de

Einleitung

Für eine auf dem edoc-Server zu veröffentlichende Dissertation müssen die Arbeit im Originalformat und ein Druckexemplar im PDF-Format vorgelegt werden. Als Originalformate werden MS Word-, Open Office- und LaTeX-Dateien akzeptiert.

Um eine PDF-Datei eines mit LaTeX erstellten Dokumentes zu erzeugen, gibt es verschiedene Wege. Der direkte Weg führt über pdfTeX.

Viele Autoren, die ihre Dissertation auf dem Dokumenten- und Publikationsserver der Humboldt-Universität veröffentlichen wollen, gehen einen anderen, längeren Weg, um die gewünschte PDF-Datei zu erhalten. Sie kompilieren ihren Quelltext mit LaTeX, um eine DVI-Datei zu erhalten. Diese DVI-Datei wird dann in PostScript konvertiert und erst in einem dritten Schritt wird aus der PostScript-Datei das PDF-Dokument.

Die Kontrolle des eingereichten Dokumentes ergibt in vielen Fällen, dass die Anforderungen des edoc-Servers an eine elektronische Publikation nicht erfüllt werden.

PDF-Funktionen bleiben ungenutzt, Schriften werden nicht richtig eingebunden oder Grafiken sind am Ende verschwunden. Die Bearbeiter sind dazu gezwungen, die Autoren zu unterrichten, wie aus den vorliegenden Quelltexten eine PDF-Datei erzeugt werden kann, die nicht nur für den Druck bestimmt ist, sondern auch den Ansprüchen an ein elektronisches Dokument genügt.

Um diese Situation zu ändern, soll es zur Pflicht werden, dass für das Erzeugen der PDF-Datei eines LaTeX-Quelltext-

es pdfTeX zum Kompilieren verwendet wird. Eine neue Vorlage für eine zur Veröffentlichung als elektronische Publikation vorgesehene Dissertation soll den Autoren dabei Hilfestellung geben und die Kontrolle der Dokumente vereinfachen.

Warum pdfTeX?

Der Einsatz von pdfTeX, anstelle des bekannten LaTeX hat zur Folge, dass Einschränkungen vorgenommen werden müssen. Zum Beispiel können Grafiken nicht im EPS-Format, das für die Archivierung abgeliefert werden muss, eingebunden werden.

Kommt pdfTeX nicht zum Einsatz, müssen die Konvertierungsschritte auch bei der Überprüfung vorgenommen werden. Das Konvertieren einer Dissertation mit vielen Grafiken von PostScript nach PDF ist leider keine Sekundensache. Die Kontrolle der Dokumente ist zeitaufwändiger.

Ob dann die von Autor und edoc-Team erstellten PDF-Dateien übereinstimmen, hängt von den verwendeten Werkzeugen und deren Konfigurationen ab. Eine mühevoll Fehlersuche beginnt, wenn die Seitenzahlen voneinander abweichen oder die Grafiken »verrutscht« sind.

Wird der Quelltext mit pdfTeX kompiliert, entfällt die Abhängigkeit von Konvertierungswerkzeugen und das PDF kann direkt erzeugt werden. Der `\pdfoutput`-Befehl ermöglicht es, weiterhin DVI-Dateien zu erstellen, wenn der Autor dies wünscht. Es weichen dann auch DVI und PDF-Version nicht mehr voneinander ab.

Für mit LaTeX verfasste Dissertationen, die auf dem edoc-Server veröffentlicht werden sollen, gelten ab April 2005 neue Vorgaben. Die wichtigsten Neuerungen sind die Forderung, das Dokument mit pdfTeX kompilieren zu können und eine Dokumentvorlage, die die Aufteilung des Dokumentes auf Filialdokumente festlegt. Beide Anforderungen wurden definiert, um PDF-Versionen der Dissertationen in hoher Qualität erzeugen zu können und die Kontrolle der Arbeiten zu vereinfachen und zu beschleunigen.

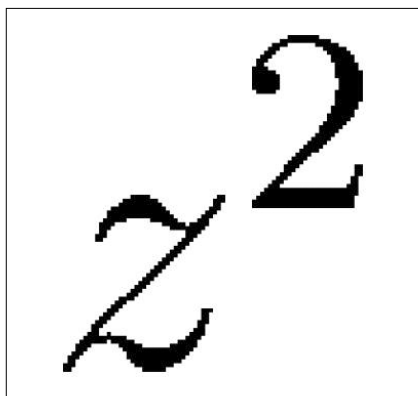


Abb. 1: Bitmap-Font.

Auch das Einbetten der Schriftarten in das zu erstellende PDF nimmt pdfTeX vor. Dafür verwendet es Map-Dateien, die die Zuordnung der LaTeX-Schriftnamen zu den PostScript-Schriftnamen beschreiben.

Das als Konvertierungswerkzeug oft verwendete dvips, das aus den DVI-Dateien PostScript-Dateien erzeugt, bindet bei entsprechend vorgenommener Konfiguration Bitmap-Fonts in die PostScript-Dateien ein, die für die Ausgabe auf dem Drucker optimiert sind. Werden aus solchen PostScript-Dateien dann PDF-Dokumente erzeugt, werden diese Bitmap-Schriften in das Dokument eingebettet. Die so eingebundenen Schriften vergrößern die Dateien und erscheinen am Bildschirm in schlechter Qualität (Vergl. Abb. 1 und 2).

Kein Vorteil ohne Nachteil

Der Nutzer von pdfTeX muss auf die Verwendung einiger LaTeX-Pakete verzichten. Zu diesen gehört PSTricks, ein Paket, mit dem Vektorgrafiken durch LaTeX-Befehle erzeugt werden können. Da die bisherigen Anforderungen an eine zu veröffentlichende Dissertation dies als »verbotene Aktion« definieren, ergibt sich somit kein Nachteil für die Autoren.

Eine zusätzliche Anforderung entsteht durch das Einbinden der Grafiken. Die für die Archivierung erforderlichen Grafiken im EPS-Format müssen nach PDF konvertiert werden, um durch pdfTeX eingebunden werden zu können. Diese Konvertierung sollte jedoch kein Problem für den LaTeX-Autor darstellen, da es hierfür frei verfügbare Werkzeuge gibt

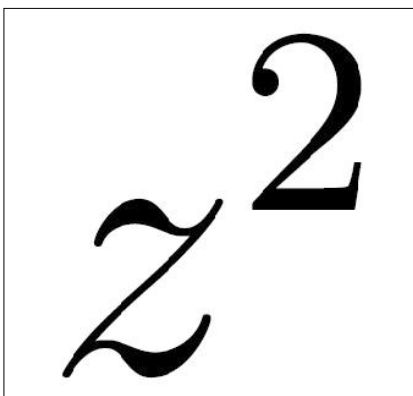


Abb. 2: PostScript-Font.

und viele LaTeX-Distributionen diese beinhalten (z. B. epstopdf).

Eine neue Dokumentvorlage?

Die neue Dokumentvorlage soll den Autoren zusätzliche Hilfestellung geben und die Kontrolle der eingereichten Dissertationen erleichtern und beschleunigen.

Diese Dokumentvorlage sieht die Aufteilung des Quelltextes auf mehrere Dateien, für die es Dateinamenskonventionen gibt, vor. Diese Aufteilung orientiert sich an der durch die *Dissertation Markup Language* (DiML) beschriebenen Struktur einer Dissertation. Die angestrebte automatisierte Kontrolle, die bei häufig auftretenden Fehlern kurze Antwortzeiten ermöglichen soll, stützt sich auf die geforderte Gliederung.

Die Verwendung des hyperref-Zusatzpaketes, die durch die neuen Vorgaben vorgeschrieben ist, soll die neue Dokumentvorlage zeigen. Hyperref erlaubt es, interaktive Dokumente zu erstellen. Es ermöglicht anklickbare Querverweise und Verzeichnisse und kann ein Inhaltsverzeichnis aus Bookmarks im Dokument ablegen.

Die korrekte Nutzung dieses Paketes stellt somit die Grundlage dar, um die Anforderungen an ein elektronisches Dokument zu erfüllen.

Informationen zur neuen Dokumentvorlage sind auf den Webseiten der Arbeitsgruppe Elektronisches Publizieren zu finden.

Unterm Strich

Die angestrebten Veränderungen bringen für Promovenden, die ihre Dissertation auf dem Dokumenten- und Publikationsserver der Humboldt Universität veröffentlichen wollen, neue Anforderungen.

Belohnt werden sie durch PDF-Dokumente, die nicht nur für die Ausgabe auf dem Drucker bestimmt sind. Auch der Aufwand für die Nachbearbeitung der Dokumente, um sie an die Anforderungen des Publikationsservers anzupassen, wird sich erheblich reduzieren, was zu kürzeren Bearbeitungszeiten bei der Annahme führen wird.

Für andere Autoren, die eine Publikation auf dem edoc-Server veröffentlichen möchten, stellen die Vorlage und die Richtlinien für eine Dissertation eine gute Grundlage dar, ein qualitativ hochwertiges PDF-Dokument zu erstellen.

Literatur

- [1] GOOSSENS, M., RAHTZ, S.: *Mit LaTeX ins Web*. Addison-Wesley, 2000.
- [2] MERZ, T., DÜMMER, O.: *Die PostScript- @ PDF-Bibel*. dpunkt-Verlag, 2002.
- [3] KIRSCH, S.: TeX akrobatisch pdfTeX erzeugt PDF ohne Umwege. *c't* 2, 2000, Seite 180.
- [4] hyperref: <http://www.tug.org/applications/hyperref/manual.html>
- [5] pdfTeX: <http://www.tug.org/applications/pdftex/index.html>
- [6] epstopdf: <ftp://ftp.dante.de/tex-archive/help/Catalogue/entries/epstopdf.html>
- [7] Dokumentvorlage: http://edoc.hu-berlin.de/e_autoren/latex/bedingung.php